

## Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

**When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.**

*This Page Blank (uspto)*

November 2005

Published\_Applications Nucleic Acid and Published\_Applications Amino Acid database searches now generate two sets of results each. The Published\_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published\_Applications\_New databases; older published applications make up the Published\_Applications\_Main databases.

Searches run against Nucleic Acid Published\_Applications produce two sets of results, with the extensions **.rnpbm** (Published\_Applications\_NA\_Main) and **.rnpbn** (Published\_Applications\_NA\_New). Searches run against Amino Acid Published\_Applications produce two sets of results, with the extensions **.rapbm** (Published\_Applications\_AA\_Main) and **.rapbn** (Published\_Applications\_AA\_New).

**This Page Blank (upio)**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using 'sw model

Run on: March 2, 2006, 18:36:26 ; Search time 123.608 Seconds  
(without alignments)  
2100.780 Million cell updates/sec

Title: US-10-693-367-2  
Perfect score: 3090  
Sequence: 1 MFGKKRVEISAPNSFEHR.....LAKGPPASIVPMQNRTR 591

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_21.\*  
1: Geneseq1980s.\*  
2: Geneseq1990s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001s.\*  
5: Geneseq2002s.\*  
6: Geneseq2003as.\*  
7: Geneseq2003bs.\*  
8: Geneseq2004s.\*  
9: Geneseq2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3090	100.0	591	2	AAY55964 Full leng
2	3090	100.0	591	3	AAY59128 Human ser
3	3090	100.0	591	7	ADC37307 Nuclear f
4	3090	100.0	591	7	ADP89973 Human can
5	3090	100.0	591	7	ADP45079 Human kin
6	3090	100.0	591	8	ADH23362 Human p21
7	3090	100.0	591	8	ABM82101 Tumour-as
8	3090	100.0	591	9	ADW71912 Human p21
9	3090	100.0	591	9	ADX06144 Cyclin-de
10	3090	100.0	591	9	ADZ64110 Human can
11	3090	100.0	620	4	ABG19308 Human hum
12	3075	99.5	588	8	ADH23359 Human p21
13	2860	92.6	593	8	ADJ96654 Human ste
14	2674.5	86.6	522	8	ABM84642 Human dia
15	2674.5	86.6	522	8	ABM84644 Human dia
16	2674.5	86.6	522	8	ABM84645 Human dia
17	2662.5	86.2	516	9	ADZ64115 Human can
18	2662	86.1	517	8	ABM84648 Human dia
19	2662	86.1	517	8	ABM84647 Human dia
20	2662	86.1	517	8	ABM84643 Human dia
21	2662	86.1	517	8	ABM84646 Human dia
22	2558	82.8	501	7	ADC37309 Nuclear f
23	2358	82.8	501	9	ADZ64111 Human can
24	2202.5	71.3	438	7	ADC37305 Nuclear f

25	2202.5	71.3	438	9	ADZ64112 Human can
26	2194.5	71.0	438	8	ADR39763 Human kin
27	2113.5	68.4	426	9	ADZ64114 Human can
28	2073	67.1	398	2	AAY55941 Human PAK
29	2069	67.0	405	9	ADZ64113 Human can
30	1748	56.6	719	7	ADJ83007 Marine ma
31	1739	56.3	719	7	ADJ83008 Human mam
32	1734	56.1	719	4	AAM38963 Human pol
33	1731	56.0	719	4	ABM85705 Novel pro
34	1731	56.0	719	7	ADC37451 Nuclear f
35	1731	56.0	719	7	ADZ64115 Human kin
36	1730	56.0	719	8	ADZ64110 Human PAK
37	1730	56.0	719	8	ADZ64111 Human p21
38	1549	50.1	457	8	ADH42207 Novel hum
39	1549	50.1	457	8	ADH42215 Novel hum
40	1549	50.1	457	8	ADH42217 Novel hum
41	1545	50.0	457	8	ADH42213 Novel hum
42	1535	49.7	457	8	ADH42211 Novel hum
43	1502	48.6	293	7	ADZ64115 Human p21
44	1497.5	48.5	639	8	ADH23364 Fruit fly
45	1494.5	48.4	635	8	ADH23360 Fruit fly

## ALIGNMENTS

### RESULT 1

AAY55964

ID AAY55964 standard; protein; 591 AA.

XX AC AAY55964;

XX DT 18-FEB-2000 (first entry)

XX DE Full length human PAK5 protein.

XX KW Antirheumatic; antiarthritic; antiinflammatory; antiallergic; osteopathic;

XX KW Antipsoriatic; antiarteriosclerotic; antidiabetic; antihypertensive;

XX KW Neuroprotective; cardioprotective; cytoprotective; cytostatic; antidiabetic;

XX KW Vulnery; SPE20; protein kinase; STLK3; STLK4; STLK5; STLK6; STLK7;

XX KW ZC1; ZC2; ZC3; ZC4; KHS2; SULUI; SULU3; GEK2; PAK4; PAK5; antagonist;

XX KW Antibody; gene therapy; rheumatoid arthritis; artherosclerosis; asthma;

XX KW Inflammatory bowel disease; Crohn's disease; osteoarthritis; psoriasis;

XX KW Rhinitis; autoimmunity; organ transplantation; multiple sclerosis;

XX KW Myocardial infarction; cardiovascular disease; stroke; renal failure;

XX KW Oxidative stress-related neurodegenerative disorder; Parkinson's disease;

XX KW Amyotrophic lateral sclerosis; Leigh syndrome; cancer; cardiomyopathy;

XX KW Ischemic disorder; inflammation; diabetes mellitus; fibrosis; mitosis;

XX KW Mesangial disorder; growth regulation; wound healing; T cell activation;

XX KW Immunosuppressant.

XX KW Homo sapiens.

XX OS WO9953036-A2.

XX PN 21-OCT-1999.

XX XX 13-APR-1999; 99WO-US008150.

XX PF 14-APR-1998; 98US-0081784P.

XX PR (SUGC-) SUGEN INC.

XX PI Plowman G, Martinez R, Whyte D;

XX DR WPI; 1999-611301/52.

XX XX N-PSDB; AA240538.

XX XX Novel kinase-related polypeptides used for the diagnosis and treatment of

XX PT kinase-related diseases and disorders.

XX PS Claim 11; Page 366-368; 387pp; English.

XX XX

CC This sequence represents a novel STE20-related protein kinase. The  
 CC invention relates to nucleic acid molecule encoding a kinase polypeptide  
 CC selected from STUK2, STUK3, STUK4, STUK5, STUK6, STUK7, ZC1, ZC2, ZC3,  
 CC ZC4, KHS2, SULJ1, SULJ3, GEK2, PAK4 and PAK5. The proteins are used to  
 CC identify agonists and antagonists, and to raise antibodies. The  
 CC polynucleotides are useful in gene therapy protocols. The polynucleotides,  
 CC polypeptides, antibodies, antagonists and agonists may be used to treat  
 CC diseases such as immune-related disorders and diseases (e.g. rheumatoid  
 CC arthritis, atherosclerosis, chronic inflammatory bowel disease (e.g.  
 CC Crohn's disease), asthma, osteoarthritis, psoriasis, atherosclerosis,  
 CC rhinitis, autoimmunity, and organ transplantation, chronic inflammatory  
 CC pelvic disease, multiple sclerosis, organ transplantation, myocardial  
 CC infarction, cardiovascular disease, stroke, renal failure, oxidative  
 CC stress-related neurodegenerative disorders (e.g. amyotrophic lateral  
 CC sclerosis, Parkinson's disease and Leigh syndrome), cancer,  
 CC cardiomyopathies, ischemic disorders, inflammatory disorders, diabetes  
 CC mellitus, fibrotic and mesangial disorders. The proteins may also be  
 CC useful for cell growth regulation (e.g. in wound healing), T cell  
 CC activation, mitosis control, and as immunosuppressants  
 XX  
 XX Sequence 591 AA;

Query Match 100.0%; Score 3090; DB 2; Length 591;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-216;  
 Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFGKKRKRVEISAPNFEHRVHTGFDQHEQKFTGLPROMQSLIEESARRPKPLVDPACIT 60  
 DB 1 MFGKKRKRVEISAPNFEHRVHTGFDQHEQKFTGLPROMQSLIEESARRPKPLVDPACIT 60

QY 61 STQGPAPKTIIVRGSKGAKDGLTLLDEFENSVTRNSLRDSDPPPPARARQENGMPEE 120  
 DB 61 STQGPAPKTIIVRGSKGAKDGLTLLDEFENSVTRNSLRDSDPPPPARARQENGMPEE 120

QY 121 PATTARGGPKAGSRFAGHSEAGGGSDRRRAGPEKPKSSRGGSGPQSSSRDKRPL 180  
 DB 121 PATTARGGPKAGSRFAGHSEAGGGSDRRRAGPEKPKSSRGGSGPQSSSRDKRPL 180

QY 181 SGPDVGTTPQAGLASGAKLAAGRPNTYPRADTDHPSRGAQGEHDVAPNGPSAGGLAIP 240  
 DB 181 SGPDVGTTPQAGLASGAKLAAGRPNTYPRADTDHPSRGAQGEHDVAPNGPSAGGLAIP 240

QY 241 QSSSSSSRPPTARCAPSPGVLPHASFPQLAPPACTPAAVPGPPGPRSPQREPORVS 300  
 DB 241 QSSSSSSRPPTARCAPSPGVLPHASFPQLAPPACTPAAVPGPPGPRSPQREPORVS 300

QY 301 HEQFRAALQLVVDGDPGRSYLDNFIKIGSGTGIVCIATVRSRSGKLVAVKMDLRKQRR 360  
 DB 301 HEQFRAALQLVVDGDPGRSYLDNFIKIGSGTGIVCIATVRSRSGKLVAVKMDLRKQRR 360

QY 361 ELLFNEVIMRDYQHENVVEMTNSYLVGDDELVWMEFLEGGALTDIVTHTRNNEBQIAAV 420  
 DB 361 ELLFNEVIMRDYQHENVVEMTNSYLVGDDELVWMEFLEGGALTDIVTHTRNNEBQIAAV 420

QY 421 CLAVLQALSVLHAQGVHRIKSDSILLTHDGRVKLSDFGCAQVSKVEPRKSLVGTTPY 480  
 DB 421 CLAVLQALSVLHAQGVHRIKSDSILLTHDGRVKLSDFGCAQVSKVEPRKSLVGTTPY 480

QY 481 WNAPELISRLPYGPEVDIWSLGTMTVEMVDGPPVFNPEPPLKAMKMRDNLPLKLNHK 540  
 DB 481 WNAPELISRLPYGPEVDIWSLGTMTVEMVDGPPVFNPEPPLKAMKMRDNLPLKLNHK 540

QY 541 VSPSLKGFLLRLIVRDPQAATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591  
 DB 541 VSPSLKGFLLRLIVRDPQAATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591

RESULT 2  
 AAY59128  
 ID AAY59128 standard; protein; 591 AA.  
 XX  
 AC AAY59128;  
 XX

DT 08-MAR-2000 (first entry)  
 XX Human serine/threonine kinase, PAK4.  
 DE PAK4; serine/threonine kinase; GTPase; intracellular signal cascade; Rac;  
 XX Cdc42H; morphogenesis; mitogenesis; JNK; p38 MAP kinase; human;  
 KW actin polymerization; filopodia; cancer; arthritis.  
 KW  
 OS Homo sapiens.  
 XX  
 XX WO9963073-A1.  
 PN 09-DEC-1999.  
 PD  
 XX 21-MAY-1999; 99WO-US011341.  
 PF  
 XX 21-MAY-1998; 98US-00082737.  
 PR  
 XX (UYCO ) UNIV COLUMBIA NEW YORK.  
 PA  
 XX Minden A;  
 PI  
 XX WPI; 2000-072881/06.  
 DR N-PSDB; AAZ40657.  
 DR  
 XX Novel mammalian nucleic acid useful for treating cancer and arthritis.  
 FT  
 XX Claim 8; Fig 1A-B; 95pp; English.  
 PS  
 XX The invention relates to an isolated mammalian nucleic acid that encodes  
 CC PAK4, a novel serine/threonine kinase or its mutant homolog. PAK4 is an  
 CC effector for the GTPases Rac and Cdc42Hs which are involved in  
 CC intracellular signal cascades, morphogenesis and mitogenesis, and  
 CC activate the JNK and p38 MAP kinase pathways. Inhibiting interaction of  
 CC PAK4 with these enzymes will thus result in inhibition of actin  
 CC polymerization and formation of filopodia. The PAK4 nucleic acid used for  
 CC recombinant production of the protein, and as a source of probes for  
 CC identifying homologous sequences and of (antisense oligonucleotides for  
 CC inhibiting PAK4 expression. The protein, or its fragments, are used to  
 CC raise specific antibodies and these are useful as ligands for therapeutic  
 CC inhibition of interaction between PAK4 and its native binding partners.  
 CC Inhibition of PAK4 activity or expression is used for treatment of cancer  
 CC and arthritis. The present sequence represents the human serine/threonine  
 CC kinase, PAK4  
 CC  
 XX Sequence 591 AA;

Query Match 100.0%; Score 3090; DB 3; Length 591;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-216;  
 Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFGKKRKRVEISAPNFEHRVHTGFDQHEQKFTGLPROMQSLIEESARRPKPLVDPACIT 60  
 DB 1 MFGKKRKRVEISAPNFEHRVHTGFDQHEQKFTGLPROMQSLIEESARRPKPLVDPACIT 60

QY 61 STQGPAPKTIIVRGSKGAKDGLTLLDEFENSVTRNSLRDSDPPPPARARQENGMPEE 120  
 DB 61 STQGPAPKTIIVRGSKGAKDGLTLLDEFENSVTRNSLRDSDPPPPARARQENGMPEE 120

QY 121 PATTARGGPKAGSRFAGHSEAGGGSDRRRAGPEKPKSSRGGSGPQSSSRDKRPL 180  
 DB 121 PATTARGGPKAGSRFAGHSEAGGGSDRRRAGPEKPKSSRGGSGPQSSSRDKRPL 180

QY 181 SGPDVGTTPQAGLASGAKLAAGRPNTYPRADTDHPSRGAQGEHDVAPNGPSAGGLAIP 240  
 DB 181 SGPDVGTTPQAGLASGAKLAAGRPNTYPRADTDHPSRGAQGEHDVAPNGPSAGGLAIP 240

QY 241 QSSSSSSRPPTARCAPSPGVLPHASFPQLAPPACTPAAVPGPPGPRSPQREPORVS 300  
 DB 241 QSSSSSSRPPTARCAPSPGVLPHASFPQLAPPACTPAAVPGPPGPRSPQREPORVS 300

QY 301 HEQFRAALQLVVDGDPGRSYLDNFIKIGSGTGIVCIATVRSRSGKLVAVKMDLRKQRR 360  
 DB 301 HEQFRAALQLVVDGDPGRSYLDNFIKIGSGTGIVCIATVRSRSGKLVAVKMDLRKQRR 360

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 18:37:25 ; Search time 49.25 Seconds  
(without alignments)  
1154.601 Million cell updates/sec

Title: US-10-693-367-2  
Perfect score: 3090  
Sequence: 1 MFGRKRVEISAPSNFEHR.....LAKAGPPASIVPLMRQNRTR 591

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 80:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1096.5	35.5	540	2 T19956	hypothetical prote
2	1090.5	35.3	542	2 T19952	hypothetical prote
3	917	29.7	544	2 A57597	beta-p21-activated
4	907	29.4	544	2 I49376	p21 activated kina
5	895.5	29.0	525	2 S58682	protein kinase, p2
6	892.5	28.9	545	2 G01773	p21-activated prot
7	892	28.9	544	2 S40482	serine/threonine-s
8	846.5	27.4	1230	2 T18256	probable serine/th
9	846.5	27.4	1230	2 T18259	serine/threonine p
10	837	27.1	939	2 S28394	probable serine/th
11	823.5	26.7	658	2 T39500	serine/threonine-s
12	814.5	26.4	658	2 S60170	protein kinase Pak
13	811.5	26.3	622	2 T15467	hypothetical prote
14	760	24.6	842	2 S60402	protein kinase CLA
15	722	23.4	378	2 T26684	hypothetical prote
16	697	22.6	589	2 T38086	serine/threonine-p
17	651.5	21.1	655	2 S51884	probable protein k
18	505	16.3	836	2 B96716	probable serine/th
19	498	16.1	471	2 T39232	probable serine th
20	494	16.0	1102	2 JC6316	probable protein k
21	478	15.5	653	2 T34356	hypothetical prote
22	476	15.4	819	2 A53714	protein kinase (EC
23	470.5	15.2	829	2 T29372	hypothetical prote
24	464	15.0	426	2 S71886	Ste20-like protein
25	463	15.0	1231	2 T18532	serine/threonine pr
26	462	15.0	1233	2 T14157	serine/threonine p
27	460	14.9	1206	2 T34021	protein kinase SK2
28	458.5	14.8	690	2 C96572	protein-F12M16.4 {
29	455.5	14.7	1080	2 S48944	hypothetical prote

30	455	14.7	1233	2 T30989	serine/threonine p
31	450.5	14.6	312	2 T38525	serine/threonine p
32	445	14.4	1001	2 T17365	hypothetical prote
33	440.5	14.3	1075	2 T27623	hypothetical prote
34	440.5	14.3	1080	2 T27622	hypothetical prote
35	437.5	14.2	1062	2 S46367	protein kinase CDC
36	428	13.9	1228	2 T18897	hypothetical prote
37	424.5	13.7	659	1 A39723	protein kinase byr
38	421	13.6	553	2 T01479	hypothetical prote
39	419	13.6	883	2 A96662	hypothetical prote
40	417.5	13.5	1174	2 T43051	protein kinase C (
41	413	13.4	561	2 T51417	protein kinase-lik
42	409.5	13.3	652	2 T39722	serine/threonine p
43	409	13.2	891	2 T40503	protein kinase kin
44	405.5	13.1	1139	1 S61918	protein kinase C (
45	404.5	13.1	1135	1 A29813	132K ninaC protein

ALIGNMENTS

RESULT 1

T19956  
hypothetical protein C45B11.lb - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 05-Oct-2004  
C;Accession: T19956  
R;McMurray, A.  
submitted to the EMBL Data Library, June 1996  
A;Reference number: Z19202  
A;Accession: T19956  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-540 <WIL>  
A;Cross-references: UNIPROT:Q5U3M1; UNIPARC:UPI000007DF1C; EMBL:Z74029; PIDN:CAA98433.1  
A;Experimental source: Clons C45B11  
C;Genetics:  
A;Gene: CESP:C45B11.lb  
A;Map position: 5  
A;Introns: 13/2; 62/3; 104/2; 231/3; 367/2; 468/3

Query Match	35.5%	Score	1096.5;	DB 2;	Length	540;
Best Local Similarity	41.7%;	Pred. No.	9.5e-32;			
Matches	251;	Conservative	67;	Mismatches	151;	Indels 133; Gaps 9;
QY	4	KRKRVEISAPSNFEHRVHTGPDHEQKFTGLTPROMQSLI--EESARRPKPLVDPAICITS	61			
Db	29	RKVKSEITPSPNFEHRIHAGFDARSCTTGLPKQWALLGPPRSISRPKPMVDPSCITP	88			
QY	62	IQFGAPKTIVRGSKGAKDGALTLLDLDEFENMVSRTSNLRRDSPPPPARARQENGMPPEP	121			
Db	89	VDVAELKTVIRGPSSRYNSPLPFGMTNSPMPSVARNSLRISATASP-----	135			
QY	122	ATTARGPGKAGSRGRFAGHSAGGSGDRRRAGPEKPKSGREGSGPQESSRDKRPLS	181			
Db	136	-----VVNVSSARHSFRPTLPPVSR-----	156			
QY	182	GPDVGTTPQAGLASGAKLAAGRPFNTYPRADTDHPSRGAQGEHDVAPNGSPSAGGLAIPQ	241			
Db	157	-----GYPF-----NDPSYAPLPL--	170			
QY	242	SSSSSRPPTRRAGAPSPGLGPH-----ASEPOLAPPACTPAAPAVPGPPG	288			
Db	171	---RNQKPM-----STTFGVEXPHQYQIITIVAPSRITTPQLQPKS-----PST	213			
QY	289	PRSPQREP---QRVSHQFRAALQLVDPGDPSPRSYLDNFIKIGEGSTGIVICIAIVRSSGK	345			
Db	214	PQAMRQPKCTEGVSDEEFENALKFVVDGTDPRSULTDYKQIGEGSTGVVEAAKYISTKQ	273			
QY	346	LVAVKMDIRKQORRELLFNEVVIWRDYOHEVVENMYSNLVCDLELVWMEFEGGALTLD	405			
Db	274	IVAVKRNLRKQORRELLFNEVSILOQYQHPNITVRFFSSHLVDDELWVWMEFEGGLTD	333			





GenCore version 5.1.7  
 Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 18:36:41 ; Search time 137.127 Seconds  
 (without alignments)  
 3040.728 Million cell updates/sec

Title: US-10-693-367-2  
 Perfect score: 3090  
 Sequence: 1 MFGKKRVRISAPNSFHR.....LAKAGFPASIVLMQRNTR 591

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : UniProt 05.80:\*  
 1: uniprot\_sprot.\*  
 2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3090	100.0	591	1 PAK4_HUMAN	O96013 homo sapien
2	2860	92.6	593	1 PAK4_MOUSE	O8btw9 mus musculu
3	2860	92.6	597	2 Q62PX0_MOUSE	O6zpx0 mus musculu
4	2853	92.3	593	2 Q80Z97_MOUSE	O80z97 mus musculu
5	2258	82.8	501	2 O8N4E1_HUMAN	O8n4e1 homo sapien
6	2202.5	71.3	438	2 O8NDE3_HUMAN	O8nde3 homo sapien
7	2202.5	71.3	467	2 Q9ULS8_HUMAN	Q9uls8 homo sapien
8	2126.5	68.8	650	2 Q5PQ75_XENLA	Q5pq75 xenopus lae
9	2124.5	68.8	650	2 Q90W62_XENLA	Q90w62 xenopus lae
10	2031	65.7	663	2 Q6GM11_BRARE	O6gm11 mus musculu
11	1967	63.7	407	2 Q8KOU2_MOUSE	O8kou2 mus musculu
12	1949	63.1	589	2 Q4STAS_TETNG	O4sta5 tetradon n
13	1748	56.6	719	2 Q8CO15_MOUSE	O8co15 mus musculu
14	1745	56.5	711	2 Q6NVJ4_BRARE	O6nvj4 brachydanio
15	1740	56.3	719	2 Q5R7U3_PONPY	Q5r7u3 pongo pygma
16	1740	56.3	719	2 Q8BVB0_MOUSE	O8bvb0 mus musculu
17	1731	56.0	719	1 PAK7_HUMAN	O9p286 homo sapien
18	1731	56.0	719	2 Q8TB93_HUMAN	O8tb93 homo sapien
19	1731	56.0	719	2 Q5W115_HUMAN	O5w115 homo sapien
20	1730	56.0	719	2 Q6RWS7_MOUSE	O6rws7 mus musculu
21	1672	54.1	908	2 Q4SCY3_TETNG	O4scy3 tetradon n
22	1507	48.8	616	2 Q7QGS5_ANOGA	O7qgs5 anopheles g
23	1501.5	48.6	639	1 PAK6_DROME	O9vxe5 drosophila
24	1476.5	47.8	681	2 Q5R8Z4_PONPY	Q5r8z4 pongo pygma
25	1473.5	47.7	681	1 PAK6_HUMAN	O9nqu5 homo sapien
26	1405	45.5	625	2 Q5SOE8_BRARE	O5sqe8 brachydanio
27	1241	40.2	240	2 O8NCH5_HUMAN	O8nch5 homo sapien
28	1160	37.5	229	2 Q9CS71_MOUSE	O9cs71 mus musculu
29	1096.5	35.5	540	2 Q9U3M1_CABEL	O9u3m1 caenorhabdi
30	1096	35.5	286	2 Q4R3T4_MACFA	O4r3t4 macaca fasc
31	1090.5	35.3	542	2 Q18637_CABEL	Q18637 caenorhabdi

32 1005 32.5 499 2 Q60MN0\_CABER  
 33 950 30.7 662 2 Q7QD76\_ANOGA  
 34 937.5 30.3 704 2 Q24190\_DROME  
 35 937.5 30.3 704 2 Q9VI13\_DROME  
 36 936 30.3 827 2 Q6IWN3\_USTMA  
 37 936 30.3 827 2 Q4PBV7\_USTMA  
 38 932.5 30.2 704 2 Q24213\_DROME  
 39 923 29.9 848 2 Q5BBL3\_EMENI  
 40 917 29.7 544 1 PAK3\_RAT  
 41 907 29.4 577 2 Q803Z0\_BRARE  
 42 907 29.4 986 2 Q4WHP3\_ASPFU  
 43 906.5 29.3 559 1 PAK3\_MOUSE  
 44 905 29.3 577 2 Q6P0I7\_BRARE  
 45 904 29.3 524 1 PAK2\_HUMAN

## ALIGNMENTS

RESULT 1  
 ID PAK4\_HUMAN STANDARD; PRT; 591 AA.  
 AC O96013; Q9BU33;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE Serine/threonine-protein kinase PAK 4 (EC 2.7.1.37) (p21-activated kinase 4) (PAK-4).  
 GN Name=PAK4;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
 OC Homo.  
 OX NCBI\_TaxID=9606;  
 RN (1) \_TaxID=9606;  
 RP NUCLEOTIDE SEQUENCE (ISOFORM 1).  
 RX MEDLINE=99043860; PubMed=9822598; DOI=10.1093/emboj/17.22.6527;  
 RA Abo A., Qu J., Cammarano M.S., Dan C., Fritsch A., Baud V., Belisle B., Minden A.;  
 RT "PAK4, a novel effector for Cdc42Hs, is implicated in the reorganization of the actin cytoskeleton and in the formation of filopodia";  
 RT EMBO J. 17:6527-6540 (1998).

RL NUCLEOTIDE SEQUENCE (ISOFORM 1).  
 RA Melnick M.B.;  
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBSJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 2).  
 RC TISSUE=Eye, Pancreas, and Placenta;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Moore T., Max S.I., Wang J., Hsieh F., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E., Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.J., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.C., Trnka J.W., Green E.D., Dickinson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC !- FUNCTION: Activates the JNK pathway. Implicated in the reorganization of the actin cytoskeleton and in the formation of



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 18:38:14 ; Search time 64.701 Seconds  
(without alignments)  
755.187 Million cell updates/sec

Title: US-10-693-367-2  
Perfect score: 3090  
Sequence: 1 MFGKKRVEISAPSNFHR.....LAKGPPASIVPLMRQNR 591

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679.residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/prodata/1/iaa/5 COMB.pep.\*
- 2: /cgn2\_6/prodata/1/iaa/6 COMB.pep.\*
- 3: /cgn2\_6/prodata/1/iaa/H COMB.pep.\*
- 4: /cgn2\_6/prodata/1/iaa/PCUS COMB.pep.\*
- 5: /cgn2\_6/prodata/1/iaa/RE COMB.pep.\*
- 6: /cgn2\_6/prodata/1/iaa/baCkfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3090	100.0	591	2	US-09-082-737-2
2	3090	100.0	591	2	US-09-688-188B-103
3	3090	100.0	591	2	US-09-718-032-2
4	3090	100.0	591	2	US-09-291-417D-103
5	3090	100.0	591	2	US-09-949-016-6665
6	3090	100.0	620	2	US-09-949-016-7206
7	2073	67.1	398	2	US-09-688-188B-30
8	2073	67.1	398	2	US-09-291-417D-30
9	1473.5	47.7	681	2	US-09-688-188B-29
10	1473.5	47.7	681	2	US-09-291-417D-29
11	1473.5	47.7	681	2	US-09-659-737A-10
12	1473.5	47.7	681	2	US-10-885-921-10
13	1473.5	47.7	694	2	US-09-949-016-10445
14	1470.5	47.6	681	2	US-09-765-815-2
15	1277	41.3	250	2	US-09-082-737-3
16	1277	41.3	250	2	US-09-718-032-3
17	1029	33.3	311	2	US-09-659-737A-8
18	1029	33.3	311	1	US-10-885-921-8
19	912	29.5	544	1	US-08-935-760-2
20	909	29.4	544	2	US-09-949-016-11562
21	907	29.4	544	2	US-09-688-188B-95
22	907	29.4	544	2	US-09-291-417D-95
23	904	29.3	524	1	US-09-538-092-1301
24	900	29.1	524	1	US-08-615-942A-2
25	900	29.1	524	2	US-09-237-325-2
26	899	29.1	544	2	US-08-559-397A-19
27	894	28.9	506	1	US-08-369-780-2

28	894	28.9	506	1	US-08-475-682-2	Sequence 2, Appli
29	894	28.9	506	1	US-08-780-833-2	Sequence 2, Appli
30	894	28.9	506	1	US-08-636-036-2	Sequence 2, Appli
31	894	28.9	506	2	US-08-918-509-2	Sequence 2, Appli
32	894	28.9	506	2	US-09-108-262-2	Sequence 2, Appli
33	894	28.9	506	2	US-09-688-188B-94	Sequence 94, Appli
34	894	28.9	506	2	US-09-291-417D-94	Sequence 94, Appli
35	893.5	28.9	551	2	US-09-949-016-10951	Sequence 10951, A
36	892.5	28.9	545	2	US-09-538-092-1297	Sequence 1297, Ap
37	891.5	28.9	545	1	US-08-935-760-4	Sequence 4, Appli
38	891.5	28.9	545	2	US-09-688-188B-93	Sequence 93, Appli
39	891.5	28.9	545	2	US-09-291-417D-93	Sequence 93, Appli
40	867	28.1	544	2	US-08-559-397A-29	Sequence 29, Appli
41	849	27.5	544	2	US-08-559-397A-30	Sequence 30, Appli
42	837	27.1	939	2	US-09-487-558B-380	Sequence 380, App
43	833.5	27.0	465	1	US-08-114-555A-2	Sequence 2, Appli
44	821	26.6	694	2	US-08-559-397A-31	Sequence 31, Appli
45	814.5	26.4	465	2	US-08-559-397A-2	Sequence 2, Appli

## ALIGNMENTS

## RESULT 1

US-09-082-737-2  
; Sequence 2, Application US/09082737  
; Patent No. 6013500  
; GENERAL INFORMATION:  
; APPLICANT: Minden, Audrey  
; TITLE OF INVENTION: PAK4; A No. 601350001 Gene Encoding A Serine/  
; NUMBER OF INVENTION: Threonine Kinase  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 11230  
; COMPUTER READABLE FORM:  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/082,737  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 0575/55311  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 591 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-082-737-2

Query Match 100.0%; Score 3090; DB 2; Length 591;

Best Local Similarity 100.0%; Pred. No. 6.6e-151;  
Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFGKKRVEISAPSNFHRVHTGDFQHEQKFTGLPQWQSLIESARRPKLVPDPACIT 60

Db 1 MFGKKRVEISAPSNFHRVHTGDFQHEQKFTGLPQWQSLIESARRPKLVPDPACIT 60

QY 61 SIQGPATKTVRSGKADGALTLLDDEFNNVTRSNLRDSDPPPARAQENGMPPEE 120

Db 61 SIQCAPKTIVRGSKGAKDGLTLLDDEFNNVTSRNSLRDSDPPPARAQENGMPDE 120  
Qy 121 PATTARGGPKAGSRGRFAGHSEAGGSGDRRRRAGPEKPKSSREGSGGPQESSSDKRPL 180  
Db 121 PATTARGGPKAGSRGRFAGHSEAGGSGDRRRRAGPEKPKSSREGSGGPQESSSDKRPL 180  
Qy 181 SGPDVGTTPQAGLAGSAGKLAAGRPFNTYPRADTDHPSRGAQGEHVDVAPNGPSAGGLAIP 240  
Db 181 SGPDVGTTPQAGLAGSAGKLAAGRPFNTYPRADTDHPSRGAQGEHVDVAPNGPSAGGLAIP 240  
Qy 241 QSSSSSSRPPTRARGAPSGVLGPHASEPQLAPACTPAAPVPGPPGPRSPQRPQVRS 300  
Db 241 QSSSSSSRPPTRARGAPSGVLGPHASEPQLAPACTPAAPVPGPPGPRSPQRPQVRS 300  
Qy 301 HEQFRAALQLVVDGPDPSYLDNFIKIGEGSTGIVCIATVRSSGKLVAVKMDLRKQRR 360  
Db 301 HEQFRAALQLVVDGPDPSYLDNFIKIGEGSTGIVCIATVRSSGKLVAVKMDLRKQRR 360  
Qy 361 ELLFNEVIMRDYQHENVVEMYSYLVGDELWVMEFLEGGALTDIVTHTRMNEEQIAAV 420  
Db 361 ELLFNEVIMRDYQHENVVEMYSYLVGDELWVMEFLEGGALTDIVTHTRMNEEQIAAV 420  
Qy 421 CLAVLQALSVLHAQGVTHRDIKSDSILLTHDGRVKLSDFGFCQAQVSKVPRKSLVGTYP 480  
Db 421 CLAVLQALSVLHAQGVTHRDIKSDSILLTHDGRVKLSDFGFCQAQVSKVPRKSLVGTYP 480  
Qy 481 WMAPELISRLPYGPEVDIWSLGIWVEMVDGPPYFNEPPLKAMKVIIRDNLPPRLKNLHK 540  
Db 481 WMAPELISRLPYGPEVDIWSLGIWVEMVDGPPYFNEPPLKAMKVIIRDNLPPRLKNLHK 540  
Qy 541 VSPSLKGFLLRDLVRDPAQRATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591  
Db 541 VSPSLKGFLLRDLVRDPAQRATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591

## RESULT 2

US-09-688-188B-103  
; Sequence 103, Application US/09688188B  
; Patent No. 6656716  
; GENERAL INFORMATION:  
; APPLICANT: PLOWMAN, GREGORY  
; APPLICANT: MARTINEZ, RICARDO  
; APPLICANT: WHYTE, DAVID  
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES  
; FILE REFERENCE: 038602/0328  
; CURRENT APPLICATION NUMBER: US/09/688,188B  
; CURRENT FILING DATE: 2000-10-16  
; PRIOR APPLICATION NUMBER: 09/291,417  
; PRIOR FILING DATE: 1998-04-14  
; PRIOR APPLICATION NUMBER: 60/081,784  
; PRIOR FILING DATE: 1998-04-14  
; NUMBER OF SEQ ID NOS: 155  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 103  
; LENGTH: 591  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-688-188B-103

Query Match 100.0%; Score 3090; DB 2; Length 591;  
Best Local Similarity 100.0%; Pred. No. 6.6e-191;  
Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MFGKKRKRVEISAPSNFHRVHTGFDQHEQKFTGLPRQWQSLIESARRPKPLVDPAKIT 60  
Db 1 MFGKKRKRVEISAPSNFHRVHTGFDQHEQKFTGLPRQWQSLIESARRPKPLVDPAKIT 60  
Qy 61 SIQCAPKTIVRGSKGAKDGLTLLDDEFNNVTSRNSLRDSDPPPARAQENGMPDE 120  
Db 61 SIQCAPKTIVRGSKGAKDGLTLLDDEFNNVTSRNSLRDSDPPPARAQENGMPDE 120  
Qy 121 PATTARGGPKAGSRGRFAGHSEAGGSGDRRRRAGPEKPKSSREGSGGPQESSSDKRPL 180  
Db 121 PATTARGGPKAGSRGRFAGHSEAGGSGDRRRRAGPEKPKSSREGSGGPQESSSDKRPL 180

Db 121 PATTARGGPKAGSRGRFAGHSEAGGSGDRRRRAGPEKPKSSREGSGGPQESSSDKRPL 180  
Qy 181 SGPDVGTTPQAGLAGSAGKLAAGRPFNTYPRADTDHPSRGAQGEHVDVAPNGPSAGGLAIP 240  
Db 181 SGPDVGTTPQAGLAGSAGKLAAGRPFNTYPRADTDHPSRGAQGEHVDVAPNGPSAGGLAIP 240  
Qy 241 QSSSSSSRPPTRARGAPSGVLGPHASEPQLAPACTPAAPVPGPPGPRSPQRPQVRS 300  
Db 241 QSSSSSSRPPTRARGAPSGVLGPHASEPQLAPACTPAAPVPGPPGPRSPQRPQVRS 300  
Qy 301 HEQFRAALQLVVDGPDPSYLDNFIKIGEGSTGIVCIATVRSSGKLVAVKMDLRKQRR 360  
Db 301 HEQFRAALQLVVDGPDPSYLDNFIKIGEGSTGIVCIATVRSSGKLVAVKMDLRKQRR 360  
Qy 361 ELLFNEVIMRDYQHENVVEMYSYLVGDELWVMEFLEGGALTDIVTHTRMNEEQIAAV 420  
Db 361 ELLFNEVIMRDYQHENVVEMYSYLVGDELWVMEFLEGGALTDIVTHTRMNEEQIAAV 420  
Qy 421 CLAVLQALSVLHAQGVTHRDIKSDSILLTHDGRVKLSDFGFCQAQVSKVPRKSLVGTYP 480  
Db 421 CLAVLQALSVLHAQGVTHRDIKSDSILLTHDGRVKLSDFGFCQAQVSKVPRKSLVGTYP 480  
Qy 481 WMAPELISRLPYGPEVDIWSLGIWVEMVDGPPYFNEPPLKAMKVIIRDNLPPRLKNLHK 540  
Db 481 WMAPELISRLPYGPEVDIWSLGIWVEMVDGPPYFNEPPLKAMKVIIRDNLPPRLKNLHK 540  
Qy 541 VSPSLKGFLLRDLVRDPAQRATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591  
Db 541 VSPSLKGFLLRDLVRDPAQRATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591

## RESULT 3

US-09-718-032-2  
; Sequence 2, Application US/09718032  
; Patent No. 6667168  
; GENERAL INFORMATION:  
; APPLICANT: The Trustees of Columbia University  
; APPLICANT: Minden, Audrey  
; TITLE OF INVENTION: PAK4, A NOVEL GENE ENCODING A SERINE/THREONINE KINASE  
; FILE REFERENCE: 575/55311-A-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/718,032  
; CURRENT FILING DATE: 2000-11-21  
; PRIOR APPLICATION NUMBER: PCT/US99/11341  
; PRIOR FILING DATE: 1999-05-21  
; PRIOR APPLICATION NUMBER: 09/082,737  
; PRIOR FILING DATE: 1998-05-21  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 591  
; TYPE: PRT  
; ORGANISM: human  
US-09-718-032-2

Query Match 100.0%; Score 3090; DB 2; Length 591;  
Best Local Similarity 100.0%; Pred. No. 6.6e-191;  
Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MFGKKRKRVEISAPSNFHRVHTGFDQHEQKFTGLPRQWQSLIESARRPKPLVDPAKIT 60  
Db 1 MFGKKRKRVEISAPSNFHRVHTGFDQHEQKFTGLPRQWQSLIESARRPKPLVDPAKIT 60  
Qy 61 SIQCAPKTIVRGSKGAKDGLTLLDDEFNNVTSRNSLRDSDPPPARAQENGMPDE 120  
Db 61 SIQCAPKTIVRGSKGAKDGLTLLDDEFNNVTSRNSLRDSDPPPARAQENGMPDE 120  
Qy 121 PATTARGGPKAGSRGRFAGHSEAGGSGDRRRRAGPEKPKSSREGSGGPQESSSDKRPL 180  
Db 121 PATTARGGPKAGSRGRFAGHSEAGGSGDRRRRAGPEKPKSSREGSGGPQESSSDKRPL 180  
Qy 181 SGPDVGTTPQAGLAGSAGKLAAGRPFNTYPRADTDHPSRGAQGEHVDVAPNGPSAGGLAIP 240  
Db 181 SGPDVGTTPQAGLAGSAGKLAAGRPFNTYPRADTDHPSRGAQGEHVDVAPNGPSAGGLAIP 240

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration, Ltd.

OM protein - protein search, using sw model

Run on: March 2, 2006, 18:40:20 ; Search time 215.348 Seconds  
(without alignments)  
1146.689 Million cell updates/sec

Title: US-10-693-367-2  
Perfect score: 3090  
Sequence: 1 MFGKKRVEISAPSNFEHR.....LAKAGPPASIVPLMRQNRTR 591

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA Main:  
1: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
3: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
4: /cgn2\_6/prodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
5: /cgn2\_6/prodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
6: /cgn2\_6/prodata/1/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3090	100.0	591	3	US-09-291-417-103
2	3090	100.0	591	4	US-10-134-102-4
3	3090	100.0	591	4	US-10-394-322A-48
4	3090	100.0	591	4	US-10-693-367-2
5	3090	100.0	591	5	US-10-725-329-103
6	3090	100.0	591	5	US-10-893-025-5
7	3090	100.0	591	5	US-10-725-121-103
8	3090	100.0	591	5	US-10-509-599-18
9	3090	100.0	620	5	US-10-450-763-49667
10	3075	99.5	588	4	US-10-134-102-1
11	2860	92.6	593	3	US-10-618-941-111
12	2073	67.1	398	3	US-09-291-417-30
13	2073	67.1	398	5	US-10-725-329-30
14	2073	67.1	398	5	US-10-725-121-30
15	1748	56.6	719	4	US-10-331-095-2
16	1731	56.0	719	4	US-10-331-095-4
17	1731	56.0	719	4	US-10-394-322A-49
18	1502	48.6	292	4	US-10-406-676-5
19	1502	48.6	292	4	US-10-406-676-6
20	1502	48.6	292	4	US-10-406-676-8
21	1502	48.6	292	4	US-10-406-676-9
22	1502	48.6	292	4	US-10-406-676-10
23	1502	48.6	292	4	US-10-406-676-11
24	1502	48.6	292	4	US-10-406-676-4
25	1497.5	48.5	639	4	US-10-134-102-6
26	1494.5	48.4	635	4	US-10-134-102-2
27	1481	47.9	292	4	US-10-406-676-7

28	1481	47.9	292	4	US-10-406-676-12	Sequence 12, Appl
29	1473.5	47.7	681	3	US-09-291-417-29	Sequence 29, Appl
30	1473.5	47.7	681	5	US-10-725-329-29	Sequence 29, Appl
31	1473.5	47.7	681	5	US-10-737-450-108	Sequence 108, Appl
32	1473.5	47.7	681	5	US-10-885-921-10	Sequence 10, Appl
33	1473.5	47.7	681	5	US-10-725-121-29	Sequence 29, Appl
34	1470.5	47.6	681	3	US-09-765-815-2	Sequence 2, Appl
35	1467	47.5	682	5	US-10-840-512-209	Sequence 209, Appl
36	1421	46.0	292	4	US-10-406-676-15	Sequence 15, Appl
37	1335.5	43.2	641	4	US-10-311-034-15	Sequence 15, Appl
38	1303	42.2	290	4	US-10-406-676-13	Sequence 13, Appl
39	1289	41.7	290	4	US-10-406-676-14	Sequence 14, Appl
40	1289	41.7	290	4	US-10-406-676-16	Sequence 16, Appl
41	1289	41.7	290	4	US-10-406-676-17	Sequence 17, Appl
42	1277	41.3	250	4	US-10-693-367-3	Sequence 3, Appl
43	1213	39.3	290	4	US-10-406-676-18	Sequence 18, Appl
44	1209	39.1	290	4	US-10-406-676-19	Sequence 19, Appl
45	1209	39.1	290	4	US-10-406-676-20	Sequence 20, Appl

## ALIGNMENTS

### RESULT 1

US-09-291-417-103  
; Sequence 103, Application US/09291417A  
; Publication No. US20030050230A1  
; GENERAL INFORMATION:  
; APPLICANT: PLOWMAN, GREGORY  
; APPLICANT: MARTINEZ, RICARDO  
; APPLICANT: WHITE, DAVID  
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES  
; FILE REFERENCE: 240/300  
; CURRENT APPLICATION NUMBER: US/09/291,417A  
; CURRENT FILING DATE: 1999-04-13  
; EARLIER APPLICATION NUMBER: US 60/081,784  
; EARLIER FILING DATE: 1998-04-14  
; NUMBER OF SEQ ID NOS: 147  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 103  
; LENGTH: 591  
; TYPE: PRT  
; ORGANISM: Full Length Mammalian (Human) PAKShu  
US-09-291-417-103

Query Match		100.0%	Score 3090;	DB 3;	Length 591;
Best Local Similarity		100.0%	Pred. No. 7.5e-161;		
Matches 591;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MFGKKRVEISAPSNFEHRVHTGFDQHEQKFTGLPRQWQSLIESARRPKPLVDPACIT	60		
Db	1	MFGKKRVEISAPSNFEHRVHTGFDQHEQKFTGLPRQWQSLIESARRPKPLVDPACIT	60		
QY	61	SIQGPATTVIRGSKAGKDGALTLLDDEFNMVTRSNLRDSDPPPARAQENGMPEE	120		
Db	61	SIQGPATTVIRGSKAGKDGALTLLDDEFNMVTRSNLRDSDPPPARAQENGMPEE	120		
QY	121	PATTARGGPKAGSRGRFAGHSEAGGSGDRRRRAGPKRPSRSGSGPQESSRDKRPL	180		
Db	121	PATTARGGPKAGSRGRFAGHSEAGGSGDRRRRAGPKRPSRSGSGPQESSRDKRPL	180		
QY	181	SGPDVGTTPQAGLASGAKLAAGRPPNTYPRADTDHPSRGAQCEPHDVPNGSAGGLAIP	240		
Db	181	SGPDVGTTPQAGLASGAKLAAGRPPNTYPRADTDHPSRGAQCEPHDVPNGSAGGLAIP	240		
QY	241	QSSSSSSRPPTTRARGAPSGVLGPHASBPQLAPPACTPAAPAVPGPPRSPQRPQVRS	300		
Db	241	QSSSSSSRPPTTRARGAPSGVLGPHASBPQLAPPACTPAAPAVPGPPRSPQRPQVRS	300		
QY	301	HEQFRAALQVLVDPGDRSYLDNFIKIGEGSTGIVCIATVRSRSGKLVAKWDLKQQR	360		
Db	301	HEQFRAALQVLVDPGDRSYLDNFIKIGEGSTGIVCIATVRSRSGKLVAKWDLKQQR	360		

QY 361 ELLFNEVIMRDYQHENVVEMNSYLVGDELWVMEFLEGGALTDIVTHTRMNEEQIAAV 420  
 Db 361 ELLFNEVIMRDYQHENVVEMNSYLVGDELWVMEFLEGGALTDIVTHTRMNEEQIAAV 420  
 QY 421 CLAVLQALSVLHAQGVIIHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRRKSILVGTPT 480  
 Db 421 CLAVLQALSVLHAQGVIIHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRRKSILVGTPT 480  
 QY 481 WMAPELISRLPYGPEVDIWSLGIWVEMVDGPPYFNEPPLKAMKWIIRDNLPPRLKNLHK 540  
 Db 481 WMAPELISRLPYGPEVDIWSLGIWVEMVDGPPYFNEPPLKAMKWIIRDNLPPRLKNLHK 540  
 QY 541 VSPSLKGFLLVDRDPAQRATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591  
 Db 541 VSPSLKGFLLVDRDPAQRATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591

RESULT 2  
 US-10-134-102-4  
 ; Sequence 4, Application US/10134102  
 ; Publication No. US20030186254A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Moritz, Michael B.  
 ; APPLICANT: Comb, Michael J.  
 ; TITLE OF INVENTION: Regulation of HIV-Tat and Nef by the Pak4 kinase and its  
 ; binding partners and methods of identifying modulators thereof.  
 ; FILE REFERENCE: CST-176 CIP  
 ; CURRENT APPLICATION NUMBER: US/10/134,102  
 ; PRIOR FILING DATE: 2002-04-29  
 ; PRIOR APPLICATION NUMBER: 09/750,457  
 ; PRIOR FILING DATE: 2000-12-28  
 ; PRIOR APPLICATION NUMBER: 60/173,939  
 ; PRIOR FILING DATE: 1999-12-30  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: Patent In Ver. 2.0  
 ; SEQ ID NO 4  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-134-102-4

Query Match 100.0%; Score 3090; DB 4; Length 591;  
 Best Local Similarity 100.0%; Pred. No. 7.5e-161;  
 Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFGKRRKVEISAPSNFHRVHTGFDQHEQKFTGLPROMQSLIESARRPKPLVDPAIT 60  
 Db 1 MFGKRRKVEISAPSNFHRVHTGFDQHEQKFTGLPROMQSLIESARRPKPLVDPAIT 60  
 QY 61 SIQCAPKTIIVRGSKGAKDGALTLLDEFENMVSRTNSLRDSDPPPPARARQENGMPPEE 120  
 Db 61 SIQCAPKTIIVRGSKGAKDGALTLLDEFENMVSRTNSLRDSDPPPPARARQENGMPPEE 120  
 QY 121 PATTARGGPKAGSRGPFAGHSEAGGSGDRRRRAGPEKPKSSRSGGPGQSSSDKRPL 180  
 Db 121 PATTARGGPKAGSRGPFAGHSEAGGSGDRRRRAGPEKPKSSRSGGPGQSSSDKRPL 180  
 QY 181 SGPDVGTTPQAGLASGAKLAAGRPFNTYPRADTDHPSRGAQGEPHDVAPNGPSAGGLAIP 240  
 Db 181 SGPDVGTTPQAGLASGAKLAAGRPFNTYPRADTDHPSRGAQGEPHDVAPNGPSAGGLAIP 240  
 QY 241 QSSSSSSRPPTTRARGAPSPVGLPHASEPQLAPACTPAAPVPGPPGPRSPQREPRQVS 300  
 Db 241 QSSSSSSRPPTTRARGAPSPVGLPHASEPQLAPACTPAAPVPGPPGPRSPQREPRQVS 300  
 QY 301 HEQFRAALQLVVDGDPSPSYLDFNFKIGEGSTGIVCIATVRSKGKLVAVKMDLKKQRR 360  
 Db 301 HEQFRAALQLVVDGDPSPSYLDFNFKIGEGSTGIVCIATVRSKGKLVAVKMDLKKQRR 360  
 QY 361 ELLFNEVIMRDYQHENVVEMNSYLVGDELWVMEFLEGGALTDIVTHTRMNEEQIAAV 420  
 Db 361 ELLFNEVIMRDYQHENVVEMNSYLVGDELWVMEFLEGGALTDIVTHTRMNEEQIAAV 420

QY 421 CLAVLQALSVLHAQGVIIHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRRKSILVGTPT 480  
 Db 421 CLAVLQALSVLHAQGVIIHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRRKSILVGTPT 480  
 QY 481 WMAPELISRLPYGPEVDIWSLGIWVEMVDGPPYFNEPPLKAMKWIIRDNLPPRLKNLHK 540  
 Db 481 WMAPELISRLPYGPEVDIWSLGIWVEMVDGPPYFNEPPLKAMKWIIRDNLPPRLKNLHK 540  
 QY 541 VSPSLKGFLLVDRDPAQRATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591  
 Db 541 VSPSLKGFLLVDRDPAQRATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591

RESULT 3  
 US-10-394-322A-48  
 ; Sequence 48, Application US/10394322A  
 ; Publication No. US20030232391A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: SUNESIS PHARMACEUTICALS, INC.  
 ; APPLICANT: Prescott, John C.  
 ; TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS  
 ; FILE REFERENCE: 39750-0006 US  
 ; CURRENT APPLICATION NUMBER: US/10/394,322A  
 ; CURRENT FILING DATE: 2003-03-20  
 ; PRIOR APPLICATION NUMBER: US 60/366,892  
 ; PRIOR FILING DATE: 2002-03-21  
 ; NUMBER OF SEQ ID NOS: 70  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 48  
 ; LENGTH: 591  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-394-322A-48

Query Match 100.0%; Score 3090; DB 4; Length 591;  
 Best Local Similarity 100.0%; Pred. No. 7.5e-161;  
 Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFGKRRKVEISAPSNFHRVHTGFDQHEQKFTGLPROMQSLIESARRPKPLVDPAIT 60  
 Db 1 MFGKRRKVEISAPSNFHRVHTGFDQHEQKFTGLPROMQSLIESARRPKPLVDPAIT 60  
 QY 61 SIQCAPKTIIVRGSKGAKDGALTLLDEFENMVSRTNSLRDSDPPPPARARQENGMPPEE 120  
 Db 61 SIQCAPKTIIVRGSKGAKDGALTLLDEFENMVSRTNSLRDSDPPPPARARQENGMPPEE 120  
 QY 121 PATTARGGPKAGSRGPFAGHSEAGGSGDRRRRAGPEKPKSSRSGGPGQSSSDKRPL 180  
 Db 121 PATTARGGPKAGSRGPFAGHSEAGGSGDRRRRAGPEKPKSSRSGGPGQSSSDKRPL 180  
 QY 181 SGPDVGTTPQAGLASGAKLAAGRPFNTYPRADTDHPSRGAQGEPHDVAPNGPSAGGLAIP 240  
 Db 181 SGPDVGTTPQAGLASGAKLAAGRPFNTYPRADTDHPSRGAQGEPHDVAPNGPSAGGLAIP 240  
 QY 241 QSSSSSSRPPTTRARGAPSPVGLPHASEPQLAPACTPAAPVPGPPGPRSPQREPRQVS 300  
 Db 241 QSSSSSSRPPTTRARGAPSPVGLPHASEPQLAPACTPAAPVPGPPGPRSPQREPRQVS 300  
 QY 301 HEQFRAALQLVVDGDPSPSYLDFNFKIGEGSTGIVCIATVRSKGKLVAVKMDLKKQRR 360  
 Db 301 HEQFRAALQLVVDGDPSPSYLDFNFKIGEGSTGIVCIATVRSKGKLVAVKMDLKKQRR 360  
 QY 361 ELLFNEVIMRDYQHENVVEMNSYLVGDELWVMEFLEGGALTDIVTHTRMNEEQIAAV 420  
 Db 361 ELLFNEVIMRDYQHENVVEMNSYLVGDELWVMEFLEGGALTDIVTHTRMNEEQIAAV 420  
 QY 421 CLAVLQALSVLHAQGVIIHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRRKSILVGTPT 480  
 Db 421 CLAVLQALSVLHAQGVIIHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRRKSILVGTPT 480  
 QY 481 WMAPELISRLPYGPEVDIWSLGIWVEMVDGPPYFNEPPLKAMKWIIRDNLPPRLKNLHK 540  
 Db 481 WMAPELISRLPYGPEVDIWSLGIWVEMVDGPPYFNEPPLKAMKWIIRDNLPPRLKNLHK 540

GenCore version 5.1.7  
 Copyright (c) 1993 - 2006 Bioceleration Ltd.  
 OM protein - protein search, using sw model  
 Run on: March 2, 2006, 18:40:55 ; Search time 25.1078 Seconds  
 (without alignments)  
 470.772 Million cell updates/sec  
 Title: US-10-693-367-2  
 Perfect score: 3090  
 Sequence: 1 MFGKKRVEISAPNFEHR.....LAKGPPASIVPLMRQNRTR 591  
 Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5  
 Searched: 135339 seqs, 20000136 residues  
 Total number of hits satisfying chosen parameters: 135339  
 Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000  
 Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : Published Applications AA New.\*  
 1: /cgn2\_6/prodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
 2: /cgn2\_6/prodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
 3: /cgn2\_6/prodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
 4: /cgn2\_6/prodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
 5: /cgn2\_6/prodata/2/pubpaa/US05\_NEW\_PUB.pep.\*  
 6: /cgn2\_6/prodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
 7: /cgn2\_6/prodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
 8: /cgn2\_6/prodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3090	100.0	591	US-10-770-726-71	Sequence 71, Appl
2	904	29.3	524	US-10-878-556A-131	Sequence 131, Appl
3	893.5	28.9	545	US-11-134-563-16	Sequence 16, Appl
4	846.5	27.4	1230	US-11-087-099-1702	Sequence 1702, Ap
5	846.5	27.4	1230	US-11-087-099-8922	Sequence 8922, Ap
6	845.5	27.4	642	US-11-087-099-7738	Sequence 7738, Ap
7	837.5	27.1	644	US-11-087-099-11826	Sequence 11826, A
8	549	17.8	855	US-11-087-099-10166	Sequence 10166, A
9	521.5	16.9	825	US-11-087-099-5581	Sequence 5581, Ap
10	516.5	16.7	825	US-11-087-099-4218	Sequence 4218, Ap
11	503.5	16.3	487	US-11-187-410-4	Sequence 4, Appli
12	495.5	16.0	547	US-10-770-726-87	Sequence 87, Appl
13	487.5	15.8	1360	US-11-241-056-14	Sequence 14, Appl
14	486	15.7	256	US-10-877-346-72	Sequence 72, Appl
15	486	15.7	256	US-11-113-424-183	Sequence 183, Appl
16	486	15.7	431	US-11-152-366-29	Sequence 29, Appl
17	484.5	15.7	443	US-11-152-366-30	Sequence 30, Appl
18	484	15.7	588	US-10-821-234-1137	Sequence 1137, Ap
19	470.5	15.2	676	US-11-087-099-8603	Sequence 8603, Ap
20	458	14.8	898	US-11-099-691-7	Sequence 7, Appli
21	458	14.8	898	US-11-124-367A-330	Sequence 330, Appl
22	458	14.8	898	US-11-124-367A-333	Sequence 333, Appl
23	451.5	14.6	275	US-11-241-056-8	Sequence 8, Appli
24	434	13.0	1236	US-11-115-086-4	Sequence 4, Appli
25	431	13.9	266	US-11-241-056-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1

US-10-770-726-71  
 ; Sequence 71, Application US/10770726  
 ; Publication No. US20050266409A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wyeth  
 ; APPLICANT: Brown, Eugene  
 ; APPLICANT: Liu, Wei  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING, PREVENTING, AND TREATING  
 ; TITLE OF INVENTION: CANCERS  
 ; FILE REFERENCE: AM101079 (031896-010000)  
 ; CURRENT APPLICATION NUMBER: US/10/770,726  
 ; CURRENT FILING DATE: 2004-02-04  
 ; NUMBER OF SEQ ID NOS: 48640  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 71  
 ; LENGTH: 591  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-10-770-726-71

Query Match 100.0%; Score 3090; DB 6; Length 591;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-175;  
 Matches 591; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

26	407	13.2	256	6	US-10-877-346-74	Sequence 74, Appl
27	398	12.9	1338	7	US-11-087-099-11043	Sequence 11043, A
28	380	12.3	876	7	US-11-087-099-10193	Sequence 10193, A
29	374	12.1	1133	6	US-10-821-234-1219	Sequence 1219, Ap
30	370.5	12.0	406	7	US-11-087-099-1555	Sequence 1555, Ap
31	369	11.9	238	7	US-11-115-086-2	Sequence 2, Appli
32	368	11.9	404	7	US-11-087-099-2344	Sequence 2344, Ap
33	368	11.9	424	7	US-11-087-099-5546	Sequence 5546, Ap
34	368	11.9	425	7	US-11-087-099-1247	Sequence 1247, Ap
35	367.5	11.9	406	7	US-11-087-099-4651	Sequence 4651, Ap
36	367	11.9	349	7	US-11-087-099-7872	Sequence 7872, Ap
37	362.5	11.7	231	7	US-11-151-601-5	Sequence 5, Appli
38	362.5	11.7	231	7	US-11-151-601-12	Sequence 12, Appli
39	362	11.7	1558	6	US-10-329-258-14	Sequence 14, Appli
40	361.5	11.7	277	7	US-11-151-601-4	Sequence 4, Appli
41	361.5	11.7	605	7	US-11-087-099-3240	Sequence 3240, Ap
42	359	11.6	1832	7	US-11-087-099-7762	Sequence 7762, Ap
43	358	11.6	278	6	US-10-055-877-149	Sequence 149, App
44	358	11.6	278	7	US-11-103-065-4	Sequence 4, Appli
45	358	11.6	278	7	US-11-151-601-10	Sequence 10, Appli



QY 361 ELLFNEVIMRDYQHENVVEMNSYLVGDELWVMEFELEGGALTDIVTHTRNNEQIAAV 420  
Db 361 ELLFNEVIMRDYQHENVVEMNSYLVGDELWVMEFELEGGALTDIVTHTRNNEQIAAV 420  
QY 421 CLAVLQALSVLHAQGVHHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRKSLVGTPY 480  
Db 421 CLAVLQALSVLHAQGVHHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRKSLVGTPY 480  
QY 481 WMAPELISRLPYGPEVDIWSLGIWIMVVDGPPYFNEPPLKAMKMIKRDNLPPRLKLNHK 540  
Db 481 WMAPELISRLPYGPEVDIWSLGIWIMVVDGPPYFNEPPLKAMKMIKRDNLPPRLKLNHK 540  
QY 541 VSPSLKGFLLRLLVDRPAQRATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591  
Db 541 VSPSLKGFLLRLLVDRPAQRATAAELLKHPFLAKAGPPASIVPLMRQNRTR 591

RESULT 2  
US-10-878-556A-131  
; Sequence 131, Application US/10878556A  
; Publication No. US20050266399A1  
; GENERAL INFORMATION:  
; APPLICANT: Hoffmann La-Roche Inc.  
; TITLE OF INVENTION: HCV regulated protein expression  
; FILE REFERENCE: 21762  
; CURRENT APPLICATION NUMBER: US/10/878,556A  
; PRIOR FILING DATE: 2004-06-28  
; NUMBER OF SEQ ID NOS: 199  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 131  
; LENGTH: 524  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: sw\_hum/pak2\_human  
; DATABASE ENTRY DATE: 1997-11-01  
US-10-878-556A-131

Query Match 29.3%; Score 904; DB 6; Length 524;  
Best Local Similarity 35.0%; Pred. No. 1.6e-46;  
Matches 211; Conservative 94; Mismatches 136; Indels 162; Gaps 13;  
QY 4 KRKRVEISAPSNFHRVHTGFDQHEQKFTGLPRQWQSLIE-----ESARRPKPLVDP 56  
Db 67 KEKERPEISPPSDFEHTIHVGFDVATGTEGTMPEQWARLLQTSNTIKLEQKKNQAVLD- 125  
QY 57 ACITSIQPGAPKTIVRGSKAGKDGALTLLLDDEFENMSVTRNSLRD--SPPPPARAQE 114  
Db 126 -----VLKPYD-----SNTVKQKYLSTFTPP-----EK 147  
QY 115 NGMPE-EPATTARGPGKAGRGFRAGHSEAGGSGDRRRRAGPKRPSRREGSGGPOES 173  
Db 148 DGLPGSTPALNAG-----TEA-----PAVVEEDDDSET 178  
QY 174 SRDKRPLSGFDVGTGPAGLAGAKLAAGRPNTYPRADTDHPSRGAQGEHVDVAPNGFS 233  
Db 179 A-----PPVIAPRPDHTKS-----IVTRSVID-FVPAVPGDSDH----- 210  
QY 234 AGGLAIPOSSSSSRPPTARCAPSPGVLGPHASEPQLAPACTPAAPAVPGPPGRSPQ 293  
Db 211 -----VDGAASKLDKQ 221  
QY 294 REPQVRSHQFRAALQLVDPDPRSYLDNFIKIGEGSTGIVCIATVRSSGKLVAKKMD 353  
Db 222 KKKPKMTDEIMEKLETVISIGDPKKYTRYEKIGQASGTVTATDVALGGEVALKQIN 281  
QY 354 LRKQRRRELLFNEVIMRDYQHENVVEMNSYLVGDELWVMEFELEGGALTDIVTHTRN 413  
Db 282 LQKQPKKELIINEILLVMKELKNPNIWFLSDYLVGDELWVMEYLAGGSLTDVVTETCMD 341  
QY 414 EQIAAVALAVLQALSVLHAQGVHHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRK 473

Db 342 EAQIAAVALAVLQALSVLHAQGVHHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRK 401  
QY 474 SLVGTPTYMADELISRLPYGPEVDIWSLGIWIMVVDGPPYFNEPPLKAMKMIKRDNLPP 533  
Db 402 TMVGTPTYMADELISRLPYGPEVDIWSLGIWIMVVDGPPYFNEPPLKAMKMIKRDNLPP 461  
QY 534 RLKLNHKVSPSLKGFLLRLLVDRPAQRATAAELLKHPFLAKAGPPASIVPL- 586  
Db 462 ELQNPKEUSPIFRDLNKLCLNDVKEKRSKAKELLOHPLKLAFLUSLTPPLIMAAKEAMK 521  
QY 587 QNR 589  
Db 522 SNR 524

RESULT 3  
US-11-134-563-16  
; Sequence 16, Application US/11134563  
; Publication No. US20050287569A1  
; GENERAL INFORMATION:  
; APPLICANT: Leong, John M.  
; TITLE OF INVENTION: ESPFU NUCLEIC ACIDS AND PROTEINS AND  
; TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: 07917-280001  
; CURRENT APPLICATION NUMBER: US/11/134,563  
; CURRENT FILING DATE: 2005-05-20  
; PRIOR APPLICATION NUMBER: US 60/573,600  
; PRIOR FILING DATE: 2004-05-20  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 545  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-134-563-16

Query Match 28.9%; Score 893.5; DB 7; Length 545;  
Best Local Similarity 36.9%; Pred. No. 6.8e-46;  
Matches 216; Conservative 82; Mismatches 166; Indels 121; Gaps 11;  
QY 4 KRKRVEISAPSNFHRVHTGFDQHEQKFTGLPRQWQSLIEESARRPKPLVDPACITSIQ 63  
Db 68 KEKERPEISLPSDFEHTIHVGFDVATGTEGTMPEQWARLLQTSN----- 111  
QY 64 PCAPKTIVRGSKAGKDGALTLLLDDEFENMSVTRNSLRDSPPPPARAQENGMPPEPAT 123  
Db 112 -----ITKSEQKKNQAVLDVLEFYNSKKT-SNSQKYS----- 144  
QY 124 TARGPGKAGSRGPAGHSEAGGSGDRRRRAGPKRPSRREGSGGPOESSRDKRPLSGP 183  
Db 145 -----FTDKSAEDYNSNALNV-----KAVSETPAVPVSEDED---DDD 181  
QY 184 DVGTPQAGLAGAKLAAGRPNT---YPRADTDHPSRGAQGEHVDVAPNGPSAGGLAIP 240  
Db 182 DDATPPP-----VIAPRPEHTKSVYTRSVI-----EPLVFTPT----- 214  
QY 241 OSSSSSSSRPPTARCAPSPGVLGPHASEPQLAPACTPAAPAVPGPPGRSPQREPORVS 300  
Db 215 RDVATSPISPTENNTTP-----PDALTENT-----EKQKKKPKMS 249  
QY 301 HEQFRAALQLVDPDPRSYLDNFIKIGEGSTGIVCIATVRSSGKLVAKKMDLRKQRR 360  
Db 250 DEEILEKLRISVSGDPKKYTRYEKIGQASGTVTAMDVATGGEVAIKOMNLOQOPKK 309  
QY 361 ELLENVIMRDYQHENVVEMNSYLVGDELWVMEFELEGGALTDIVTHTRNNEQIAAV 420  
Db 310 ELTINEILLVMKELKNPNIWFLSDYLVGDELWVMEYLAGGSLTDVVTETCMDEQIAAV 369  
QY 421 CLAVLQALSVLHAQGVHHRDIKSDSILLTHDGRVKLSDFGCAQVSKVPRKSLVGTPY 480  
Db 370 CRECLQALEFLHSNQVHHRDIKSDSILLMDGSGVKLTDFGCAQVITPQESKSTWGTPT 429